

Analysis of Latency and Throughput of 2D Torus Topology using Modified XY Routing Algorithm

Ms. Priyanka N. Chopkar¹ and Dr. Mahendra A. Gaikwad²

¹PG Student, M-Tech (Electronics), B. D. College of Engineering Sewagram (Wardha), Maharashtra, India
priyachopkar@gmail.com

²Professor & Dean R & D, B. D. College of Engineering Sewagram (Wardha), Maharashtra, India
gaikma@rediffmail.com

Abstract - Network-on-Chip (NoC) is a packet switched on-chip communication network designed using a layered methodology i.e. “routes packets, not wires”. It is an efficient on chip communication architecture for System on Chip (SoC) architectures. NoCs solved the disadvantages of SoCs and are scalable. Each route in NoC includes some routers and it takes a few clock periods by passing a router.

When the network is in congestion, the package transmission will produce much more time delay. So adopting an appropriate routing algorithm to get the balance between the time delay and throughput rate becomes the key problem. In this paper we tried to solve that problem using torus topology with our modified XY routing algorithm.

We used NIRGAM simulator for analysis of latency and throughput of modified XY routing algorithm for 2D torus topology. 3x3 network size used for analyze the performance. We consider all tiles as source, all tiles as destination and vary the packet size & traffic used is Constant Bit Rate (CBR) random traffic. It is found that packet size increases then latency/packet increases and throughput (in Gbps) also increases but latency/flit decreases.

Keywords – 2D torus topology, modified XY routing algorithm, Network on Chip (NoC), NIRGAM simulator.

I. INTRODUCTION

NoC is an approach to design the communication subsystem between IP cores in a SoC. NoCs are expected to overcome scalability and performance limitations of Point-to-Point (P2P) and bus-based communication systems.

NoCs use packets to route data from the source to the destination Processing Elements (PE) via a network fabric that consists of switches (routers) and interconnection links (wires). The wires in the links of the NoC are shared by many signals. A high level of parallelism is achieved, because all links in the NoC can operate simultaneously on different data packets. Therefore, as the complexity of integrated systems keeps growing, a NoC provides enhanced performance (such as throughput) and scalability in comparison with communication architectures (e.g., dedicated point-to-point signal wires, shared buses, or segmented buses with bridges).

NoCs are an attempt to scale down the concepts of large scale networks, and apply them to the embedded system-on-chip (SoC) domain. The main difference between SoC and NoC is the sharing of links. The most important features that distinguish NoC architectures are network topology and

routing algorithms. Communication performance of a NoC depends heavily on the routing algorithm used.

Routing is one of the most crucial key factors which will decide over the success of NoC architecture based systems or their failure. There are different routing algorithms such as XY routing algorithm, odd even routing algorithm and source routing algorithm etc. Modified XY routing algorithm is simple and widely used. It is static, deterministic and deadlock free routing algorithm. In deterministic routing algorithm the path between source to destination is fixed.

In modified XY routing algorithm [1], a packet must always be routed along horizontal or X axis until it reaches the same column as that of destination then it should be routed along vertical or Y axis and towards the location of destination resource.

Routing is in deadlock when two packets are waiting each other to be routed forward. Both of the packets are waiting each other to release the resources. Routers do not release the resources before they get the new resources and so the routing is locked.

Topology is a very important feature in the design of NoC because design of a router depends upon it. There are different types of topologies such as mesh, torus, folded torus, star, octagon, tree, etc. Among this topology, torus topology was proposed to reduce the latency of mesh and keep its simplicity.

The main problem with the mesh topology is its long diameter that has negative effect on communication latency. The only difference between torus and mesh topologies is that the switches on the edges are connected to the switches on the opposite edges through wrap-around channels. The performance parameters are latency and throughput. Latency is defined as time taken to deliver a packet from source to destination. Throughput is defined as a fraction of packets delivered from sources to destinations in a given amount of time.

II. RELATED WORK

Mohsen Nickray et al. [2] have described an adaptive routing algorithm which is based on deterministic XY routing algorithm. In their model a switch is a context-aware agent and a network is a society of context-aware agents which are ever learning and adapting to distribute the congestion uniformly and isolate the malformed switches (agents). In conventional XY routing, first, the load in the center of a

network is much higher rather than total average and this leads to hot spot in the center of network. And second, a malfunction in switches could make part of network out of access. But in their proposed routing, first, all agents are aware from their neighbor's congestion and collaboratively try to route their input packets through less congested route, according to their experiences learned before and second, when a new malfunction takes place in the network, all agents collaborate each other to recognize the malformed agent and learn the best route. The main objective of their routing algorithm is to distribute network load uniformly throughout the network. They developed a NOC environment using SystemC.

Manas Kumar Puthal et al. [3] have described a new hierarchical cluster based adaptive routing called „C-Routing in 2D mesh NoC. The solution reduces routing table size and provides deadlock freedom without use of virtual channels while ensuring live lock free routing. Routers in this method use intelligent routing to route information between the processing elements ensuring the correctness, deadlock freeness, and congestion handling. This method has been evaluated against other adaptive algorithms such as PROM, and Q-Routing etc. C-routing uses adaptively to avoid congestion by uniform distribution of traffic among the cores by sending flits over two different paths to the destination. Proposed technique achieves two objectives, as inferred from the results, reduction in area and higher throughput. These benefits are achieved at a marginal increase in power consumption and latency while preserving deadlock freedom with no extra virtual channels. Fault tolerance is another major issue in NoC design.

Jin-xiang Wang et al. [4] have described a new fault model, defines separately node-fault and link-fault, reduces situations classified as node-fault effectively and consequently improves the performance of the network. By defining some new paths to substitute failure paths, data packets can be routed along the new paths which are formed by the neighbor nodes of node-fault or link-fault. A fault-tolerant wormhole router based on XY routing algorithm is designed according to the solution. The evaluation results show that network performance can be improved by 15% when link-fault occurs in the network.

Ruizhe Wu et al. [5] have been described a new on-chip communication system & dubbed Wireless Network on- Chip (WNoC). This work centers on the design of a high-efficient, low-cost, deadlock-free routing scheme for domain-specific irregular mesh WNoCs. A distributed minimal table based routing scheme is designed to facilitate segmented XY-routing. Deadlock-free data transmission is achieved by implementing a new turn class's based buffer ordering scheme.

Hamed S. Kia and Cristinel Ababei [6] have described algorithm is based on the ball-and-string model and employs a distributed approach based on partitioning of the regular NoC architecture into regions controlled by local monitoring units. Each local monitoring unit runs a shortest path computation procedure to identify the best routing path so that highly congested routers and faulty links are avoided while

latency is improved. To dynamically react to continuously changing traffic conditions, the shortest path computation procedure is invoked periodically. Because this procedure is based on the ball-and-string model, the hardware overhead and computational times are minimal. Experimental results based on an actual Verilog implementation demonstrate that the proposed adaptive routing algorithm improves significantly the network throughput compared to traditional XY routing and DyXY adaptive algorithms.

A.H. Borhani et al. [7] described a new fault tolerant routing algorithm, which is based on dimension order routing, is proposed for k-ary 2-cubes. Packets are sent to their destination through XY routing algorithm and if this transmission is not possible, YX routing algorithm is applied. The result shows that these method is preferred, especially in the environments where the fault probability is low and the message generation rate is high.

Yonghui Li & HuaxiGu [8] described a new model called the XY-turn model for designing partially adaptive or deterministic routing algorithms for honeycomb networks on chip without adding virtual channels. The model prohibits two turns in opposite directions at some particular nodes so that the deadlock can be avoided. The restricted turns result in simplified router architectures especially for photonic NoCs. They have implemented a deterministic routing algorithm is carried out to test the performance of the algorithm in end to end delay and throughput under the conditions that messages length are 128 bytes, 256 bytes and 512 bytes respectively, as well as the network size varies from 36 cores to 64 cores and 98 cores. The results deterministic that the honeycomb NoCs would benefits from the XY-turn model in terms of communication efficiency.

Masood Dehyadgari et al. [9] have described a pseudo adaptive routing which is an extension of classic XY routing. They consider mesh topology for evaluating proposed routing. Their switches use Pseudo adaptive XY routing algorithm. The load in the center of a network in ordinary XY routing is much higher rather than total average. This extra load on the center of mesh can cause spot hot. The main objective of their routing algorithm is to distribute network load. One of the advantages of distributing network load is balanced temperature on the mesh. Their routing algorithm has two modes that is deterministic and adaptive Packets are routed with classic XY routing (deterministic mode), when congestion in the network is low. When congestion is high, packets will be routed through less congested route adaptively (adaptive mode).

Mehrdad Seyrafi et al. [10] has been described a new fault tolerant routing algorithm with minimum hardware requirements and extremely high fault tolerance for 2D-mesh based NoCs is proposed. The LCFT (Low Cost Fault Tolerant) algorithm, removes the main limitations (forbidden turns) of the famous XY. So not only many new routes will be added to the list of selectable paths as well as deadlock freedom, but also it creates high level of fault tolerance. All these things are yielded only by the cost of adding one more virtual channel (for a total of two). Results show that LCFT algorithm

can work well under almost bad conditions of faults in comparison with the Modified-XY, Modified-YX algorithms. Yang Quansheng & Wu Zhekai [11] have described an improved topology called Tmesh, which is based on standard mesh network by inserting four long links to connect the vertices to reduce the communication delay between some remote nodes. They also present a deadlock-free routing algorithm for Tmesh named TXY algorithm. The results of this algorithm show a certain reduction in the average packet delay and routing hops. When the network has 64 nodes, the average delay and routing hops of Tmesh are 2.92% and 3.53% lower than those of mesh respectively.

Shu Yan Jiang et al. [12] describe an online detection method of interconnection for 2D torus structure of NoC system. This method can detect the data errors during transmission and identify the error results from the routing switch failure or the data transmission interconnection line failure. They design a sub-router based on the wormhole exchange using E-cube routing algorithm and a check module which is suitable for the original routing node functions and work feature. They simulate the method by Verilog HDL and quartus II software. The experiment results show that the method can detect data errors caused by the router failure or interconnect failure and can locate the fault.

Xiaohang Wang et al. [13] described a simple, yet efficient hardware based multicasting scheme is proposed for irregular mesh based NoC. First, an irregular oriented multicast strategy is proposed. Following this strategy, an irregular oriented multicast routing algorithm can be designed based on any regular mesh based multicast routing algorithm. One such algorithm, namely, Alternative XY (AL+XY), is proposed based on XY routing. Experimental results show that AL+XY achieve significant reduction in power consumption and packet latency compared with existing solutions. AL+XY saves 29% power consumption than that of multiple unicast. In terms of average packet latency, when injection rate is high (e.g. near 0.15), the latency of AL+XY is only 50%.

Slavisa Jovanovic et al. [14] has been described a new deadlock free fault tolerant adaptive routing algorithm for 2D mesh NoC interconnection. The main contribution of this routing algorithm is that it allows both, routing of messages in the networks incorporating the regions not necessarily rectangular and routing to all nodes which are not completely blocked by faulty nodes. The proposed routing algorithm is based on a modified turn model and well known XY algorithm. The basic principle of this routing algorithm, prove its deadlock freeness, its feasibility and efficiency through the simulation results.

Xiaoqiang Yang et al. [15] described node coding and routing methods are important to the design of NoC. By the combination of network topology with corresponding, a two dimensional code based on Johnson code in Torus topology is proposed. The node coding implies the relation between neighboring nodes and has a good scalable characteristic. The two methods for code compressing are also presented to reduce the storage space of node address and increase the utilization rate of channel bandwidth. The improved

algorithm for XY routing based on the code is presented and node structure is designed. The experimental results show combination of the code can simplify the routing algorithm in the implementation of NoC, decrease silicon resource consumption and greatly improve communication performance.

III. PROBLEM IDENTIFICATION

We observed in related work on design of NoC architecture that maximum paper used deterministic XY routing or adaptive XY routing with 2D-Mesh topology to reduce the latency & improve the throughput. Up to yet no one used 2D torus topology to improve the throughput.

In NIRGAM simulator the existing XY routing algorithm only works on 2D Mesh topology, it cannot work on 2D Torus topology. We propose and execute the application of XY routing algorithm is modified for 2D torus topology.

To simulate modified XY – routing algorithm for two dimensional torus topology of Network on Chip architecture for constant bit rate random traffic in NIRGAM simulator to reduce the average latency per flit, increase average latency per packet & increase average throughput per channel basis only for 3x3 network size.

IV. MODIFIED XY ROUTING ALGORITHM FOR 2D MESH AND 2D TORUS TOPOLOGY

Modified XY routing is a dimension order routing which routes packets first in x- or horizontal direction to the correct column and then in y- or vertical direction to the receiver. Modified XY routing suits well on a network using mesh or torus topology. Addresses of the routers are their xy-coordinates. Modified XY routing never runs into deadlock. The fig. 1 and fig. 2 shows the comparison of 2D mesh and 2D torus topology. The mesh and torus topology are the same with only one difference, in torus the corner and border nodes are directly connected. Following figures represents the popular mesh and torus topology. Here the rectangular boxes represent the routers while arrow represents a bi-directional link.

In order to describe working of modified XY routing algorithm, consider an example of modified XY routing algorithm for 2D mesh and 2D torus in 3x3 network size as shown in fig. 1 and fig. 2. Assume that a tile-3 acts as source and its co-ordinates are (1, 0), wants to send a packet to the destination tile-8 and its co-ordinates are (2, 2).

The route of packet for mesh topology is tile-0 to tile-4, tile-4 to tile-5, and tile-5 to tile-8, the three hops are required for moving packet from source to destination. Similarly, for torus topology the packet route is tile-3 to tile-5 and, tile-5 to tile-8, the two hops are required for moving packet from source to destination. The number of hops increases the overall latency increases. The overall latency is high in 2D mesh topology as compared to 2D torus topology.

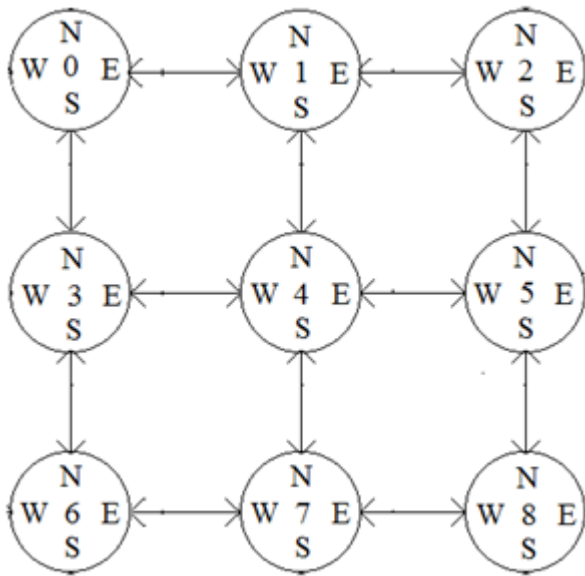


Figure 1. 2D Mesh topology

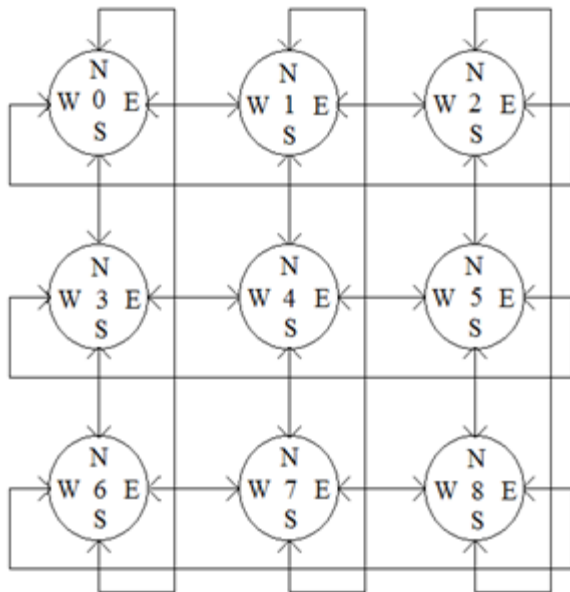


Figure 2. 2D Torus topology

V. MODIFIED XY ROUTING ALGORITHM FOR 2D TORUS TOPOLOGY

Success & failure of NOCs architecture is decided by routing algorithm. It determines the path selected by a packet to reach its destination. It must decide within each intermediate router which output channels are to be selected for incoming packets. Routing on NoC is quite similar to routing on any network. A routing algorithm determines how the data is routed from source to destination.

Modified XY routing algorithm routes packets every time from source to destination along a fixed path. It is used in both regular and irregular networks. In congestion free networks this algorithm is reliable and has low latency. It suits well on real time systems because packets always reach the destination in correct order and so a reordering is not necessary. It has better latency at low traffic.

Dimension order routing (DOR) is a typical minimal turn algorithm. The algorithm determines to what direction packets are routed during every stage of the routing. Modified XY routing is a deterministic and static routing which routes packets first in X or horizontal direction to the correct column and then in Y or vertical direction to the destination.

Fig. 2 shows the 3x3 network size by using torus topology, considered all tiles i. e. tile-0 to tile-8 as source and all tiles i.e. tile-0 to tile-8 considered as destination, that means all tiles sends packets and all tiles received packets.

VI. FLOWCHART OF MODIFIED XY ROUTING ALGORITHM FOR 3X3 2D TORUS TOPOLOGY

A Torus network is an improved version of basic mesh network. A simple torus network is a mesh in which the heads of the columns are connected to the tails of the columns and the left sides of the rows are connected to the right sides of the rows. Following fig. 3 shows flowchart of modified XY routing algorithm for 3x3 2D torus topology.

In order to demonstrate working of modified XY routing algorithm, consider an example of modified XY routing algorithm for 2D torus in 3x3 network size as shown in fig. 4.

Assume that a tile-0 acts as source and its co-ordinates are (0, 0), wants to send a packet to the destination tile-5 and its co-ordinates are (1, 2). Packet route of source to destination is indicated by black arrow in the fig. 4.

Developed algorithm first calculates difference between destination y-coordinates and source y-coordinates. In this case the y-distance is '2' thus the packet travel in west direction i.e along x-axis. After that it calculates difference between destination x-coordinates and source x-coordinates, the difference is '1' then packet move in south direction i.e along y-axis.

In this way the modified XY routing algorithm run on 2D torus topology.

VII. SIMULATION RESULTS AND ANALYSIS

The simulation has been performed on NIRGAM simulator, a simulator for NoC Interconnect Routing and Application Modeling version 2.1. NIRGAM is an extensible and modular systemC based simulator [16].

All tiles used as source and all tiles used as destinations. All tiles send packets and all tiles receive packets by using Constant Bit Rate (CBR) traffic generator. The packet size is vary from 8 bytes to 128 bytes with a random destination mode. The load percentage is 50% which means that 50% of maximum bandwidth is used. The simulation runs 1000 clock cycles and clock frequency is 1GHz. The synthetic traffic generators generate traffic in 300 clock cycles with warm up period of 5 clock cycles. The flit interval between successive flits is 2 clock cycles.

Simulation performance of routing algorithm is measured on a per-channel basis. The performance metrics are average latency (in clock cycles) per flit, average latency (in clock cycles) per packet and average throughput (in Gbps) for each channel.

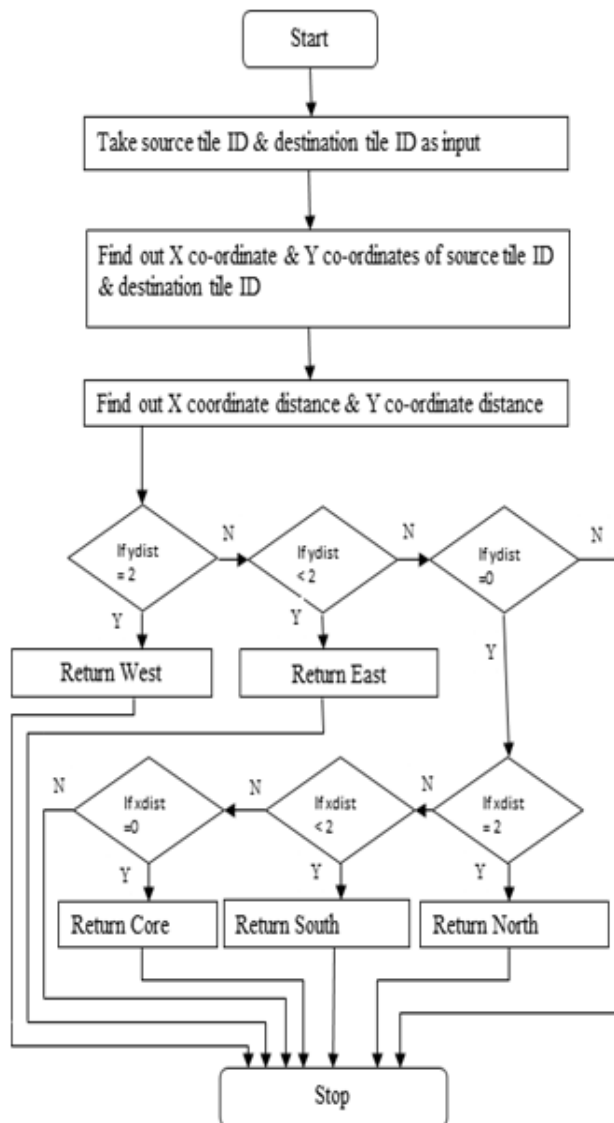


Figure 3. Flowchart of modified XY routing algorithm for 2D torus topology

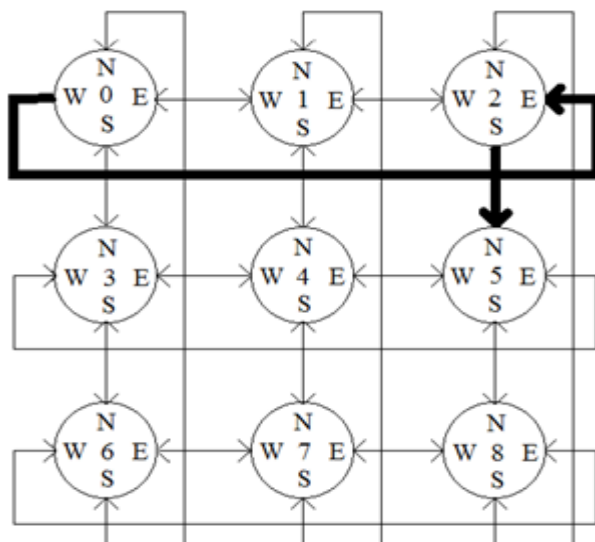


Figure 4. Example of modified XY routing algorithm for 2D torus in 3x3 network size

Table I shows simulation results of modified XY routing algorithm for 2D torus topology of NoC architecture for 3x3 network size. Packet size increases from 8 byte to 128 bytes and measured average latency per flit (in clock cycles), average latency per packet (in clock cycles) and average throughput (in Gbps) per channel.

We can conclude that in fig. 5, if packet size increases then average latency per packets increase.

If packet size increases then total no. of flit generated increases and average latency per packet increases which is shown in fig. 6. The flit size is constant i. e. 5 bytes (1 byte for head payload and 4 bytes for data payload) and flit generated increases. Because packet size increases the number of flit generated in a packet increases.

It is observed from fig. 7, total number of packet is large then latency per flit large because time taken to deliver a packet from source to destination is large. If packet size (in bytes) increases then average throughput (in Gbps) per channel also increases which is observed in fig. 8.

In fig. 9 concluded that the average throughput (in Gbps) per channel increases then average latency per packet also increases.

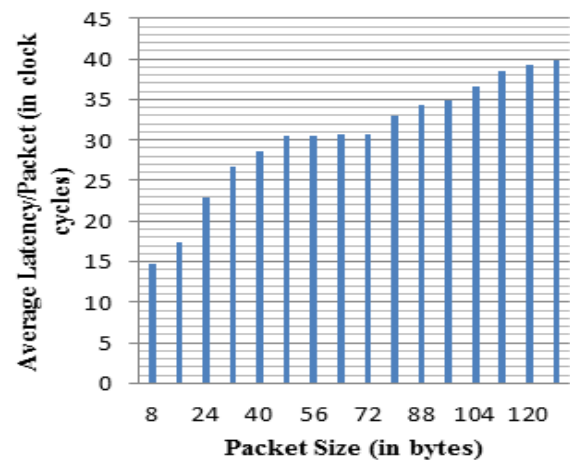


Figure 5. Packet size (in bytes) versus average latency per packet (in clock cycles)

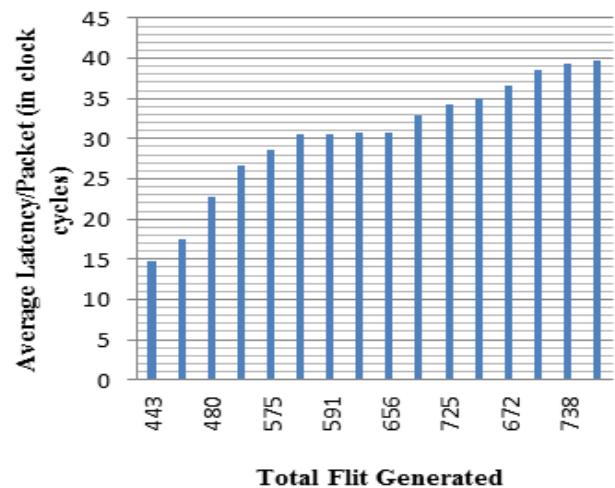


Figure 6. Total flit generated versus average latency per packet (in clock cycles)

TABLE I. SIMULATION RESULTS OF MODIFIED XY ROUTING FOR 2D TORUS TOPOLOGY IN 3x3 NETWORK SIZE

Packet Size (in bytes)	Total Flit Generated	Total Flit Received	Total No. of Packet Generated	Total No. of Packets Received	Total Packet Loss	Average Latency per flit (in Clock cycles)	Average Latency per packet (in Clock cycles)	Average Throughput (in Gbps)
8	443	195	147	65	82	4.9021	14.7592	8.6095
16	462	220	92	44	48	3.4593	17.4534	8.9
24	480	245	68	35	33	3.5038	22.8161	9.9512
32	562	288	62	32	30	2.9296	26.7428	10.1806
40	575	308	52	28	24	2.5492	28.5377	11.5718
48	579	286	44	22	22	2.3042	30.5749	12.7597
56	591	300	39	20	19	1.99305	30.4917	12.8768
64	663	425	39	25	20	1.7364	30.7042	13.9108
72	656	380	34	20	14	1.4883	30.7049	13.8341
80	671	357	32	17	15	1.3676	32.9186	14.6887
88	725	460	31	20	11	1.333	34.3359	15.4938
96	675	450	27	18	9	1.2998	34.9846	15.1962
104	672	351	25	13	12	1.07157	36.6	15.2526
112	748	427	25	14	11	1.17859	38.5294	15.1196
120	738	465	24	15	9	1.11835	39.2667	15.3165
128	781	429	23	13	10	1.101943	39.81	15.7006

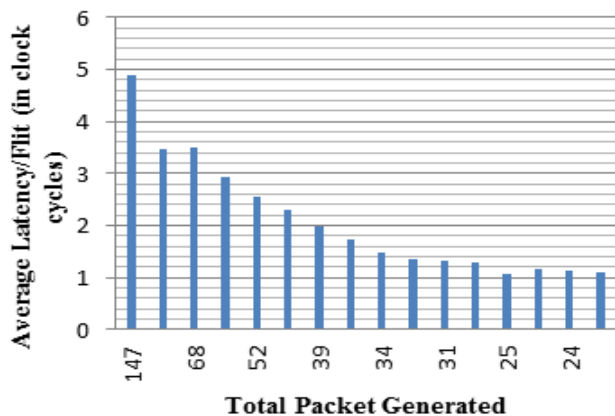


Figure 7. Total no. of packets generated (in bytes) versus average latency per flit (in clock cycles)

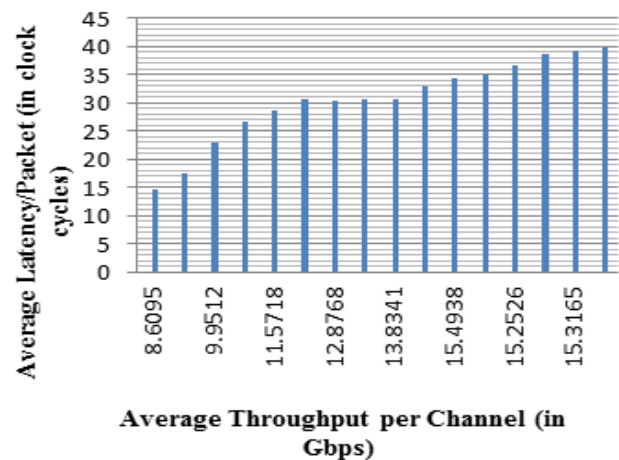


Figure 9. Average throughput (in Gbps) versus average latency per packet (in clock cycles)

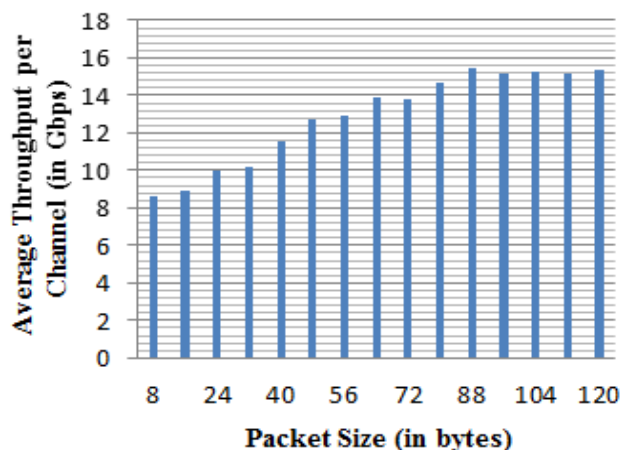


Figure 8. Packet size (in bytes) versus average throughput (in Gbps)

CONCLUSION

In this paper, we designed modified XY routing algorithm on regular 2D torus topology of NoC architecture. NIRGAM simulator is used for analyze the performance of 3x3 network size. Parameters consider are, all tiles as source and all tiles as destination, that means all tiles send packets and all tiles receive packets by using CBR random traffic generator. The packet sizes vary from 8 bytes to 128 bytes and all other parameters are constant. The performance parameters are latency and throughput. Simulation performance of routing algorithm is measured on a per-channel basis. It is observed from the result that if packet size increases then average latency per packet (in clock cycles) increases and average

throughput (in Gbps) also increases but average latency per flit (in clock cycles) decreases. All tiles act as source and all tiles act as destination, then congestion in the network therefore average latency per packet increases.

REFERENCES

- [1] Priyanka N. Chopkar and Mahendra A. Gaikwad, "Review of XY routing algorithm for 2D torus topology of NoC architecture", 2013, International Journal of Computer Applications (0975 – 8887), pp. 22-26.
- [2] Mohsen Nickray, Masood Dehyadgari & Ali Afzali- kusha, "Adaptive Routing Using Context-Aware Agents for Networks on Chips ", IEEE, 2009, pp. 1-6.
- [3] Manas Kumar Puthal, Virendra Singh, M. S. Gaur, Vijay Laxmi, "C-Routing: An Adaptive Hierarchical NoC Routing Methodology", 2011, IEEE/IFIP 19th International Conference on VLSI and System on Chip, pp. 392 – 397.
- [4] Jin-xiang Wang, Fang-fa Fu, Tian-Sheng Zhang & Yu-Ping Chen, "A Small-Granularity Solution on Fault Tolerant in 2D-Mesh Network-on-Chip", IEEE, 2010, pp. 382-384.
- [5] Ruizhe Wu, Yi Wang & Dan Zhao, "A Low Cost Deadlock-free Design of Minimal-Table Rerouted XY-Routing for Irregular Wireless NOCs", 2010, Fourth ACM/IEEE International Symposium on Networks-on-Chip, pp. 199 – 206.
- [6] Hamed S. Kia and Cristinel Ababei, "A new Fault tolerant and Congestion-aware Adaptive Routing Algorithm for Regular Networks-on-Chip", IEEE, 2011, pp. 2465-2472.
- [7] A.H. Borhani, A. Movaghar, R. G. Cole, "A New Deterministic Fault Tolerant Wormhole Routing Strategy For k-ary 2-cubes", IEEE, 2010, pp. 1 – 7.
- [8] Yonghui Li, HuaxiGu, "XY-Turn Model for Deadlock Free Routing in Honeycomb Networks on Chip", IEEE, 2009, pp. 900 – 903.
- [9] Masood Dehyadgari, Mohsen Nickray, Ali Afzali-kusha, Zainalabein Navabi, "Evaluation of Pseudo Adaptive XY Routing Using an Object Oriented Model for NOC", IEEE, 2005, pp. 204-206.
- [10] Mehrdad Seyrafi, Arghavan Asad, Amir Ehsani Zonouz, Reza Berangi, Mahmood Fathy, Mohsen Soryani, "A New Low Cost Fault Tolerant Solution for Mesh based NoCs", 2010 International Conference on Electronics and Information Engineering (ICEIE 2010), pp. V2-207 - V2-213.
- [11] Yang Quansheng, Wu Zhekai, "An Improved Mesh Topology and Its Routing Algorithm for NoC", IEEE, 2010, pp. 1-4.
- [12] Shu Yan Jiang, Hao Liang, Shuo Li, Yong Le Xie, "A Test Method of Interconnection Online Detection of NoC Based on 2D Torus Topology", IEEE, 2011, pp. 183 – 187.
- [13] Xiaohang Wang, Mei Yang, Yingtao Jiang, Peng Liu, "Efficient Multicasting Scheme for Irregular Mesh-Based NoCs", IEEE, 2010, pp. 384 – 387.
- [14] Slavisa Jovanovic, Camel Tanougast, Serge Weber, Christophe Bobda, "A New Deadlock-Free Fault-Tolerant Routing Algorithm For NoC Interconnections", IEEE, 2009, pp. 326 – 331.
- [15] Xiaoqiang Yang, Huimin Du, Jungang Han, "A Node Coding and the Improved Routing Algorithm in Torus Topology", IEEE, 2008, pp. 443 – 447.
- [16] Wang Zhang, Wuchen Wu, Lei Zuo, Xiaohong Peng, "The Buffer Depth Analysis of 2-Dimension Mesh Topology Network-on-Chip With Odd-Even Routing Algorithm", IEEE, December 2009.